

REMARKS

This paper is being filed as a response to the Office Action of March 3, 2010. Reconsideration is respectfully requested in view of these clarifying remarks. No amendments to the claim are being submitted. If the Examiner feels that a new search is needed, Applicant submits that a new non-final Office Action should be issued.

Rejections under 35 USC § 103(a)

Claims 1-19, 21-23, and 25-37 have been rejected under 35 U.S.C. 103(a) as being unpatentable over Sutton et al. (U.S. Patent No. 6,539,354), in view of Dietz (U.S. Patent No. 6,385,586). This rejection is respectfully traversed. Applicants respectfully request reconsideration of these rejections in light of the arguments contained herein.

1. The combination of the prior art does not teach altering the content data including an applied expression that **does not perform language translation.**

Claim 1 specifies altering the content data that is to be output by the second computer in accordance with the content data output characteristics specified through the first computer, the output characteristics identifying an expression to be applied to the content data. Further, claim 1 specifies that the altering includes converting an audio component of the content data to text data through a voice recognition process, the text data being processed into converted text data, and the converted text data being synthesized into audio data that includes the applied expression that does not perform language translation.

Thus, the altering includes the following 3 operations:

1. Converting an audio component of the content data to text data,

2. Processing the text data into converted text data, and

3. Synthesizing the converted text data into audio data that includes the applied expression that does not perform language translation (emphasis added).

Applicant respectfully asserts that the Office's rejection is inconsistent. First, the Office has asserted that Sutton teaches that "the altering includes converting an input component of the content data (text or voice) to multimedia output format (audio and visual speech), which is synthesized audio data that includes the applied expression that does not perform language translation" (page 3, lines 11-14). Unfortunately, the Office has indicated what Sutton teaches and not what feature of the claims Sutton teaches, (in the rest of the rejection of claim 1, the Office did indicate how the prior art supposedly teaches claim 1). It is not clear to which of the 3 operations mentioned above the Office is referring to in the rejection. In any case, it is clear that this rejection is not referring to operation 2 (processing the text data into converted text data) because the Office refers to translation into audio or visual speech and not into text. Adding to the inconsistencies is the admission by the Office that "Sutton et al. do not teach that ... the altering includes converting an audio component of the content data to text data, the text data being processed into converted text data, and the converted text data being synthesized into audio data" (page 3, last para., emphasis added). This is a contradiction by the Office, as the Office has admitted that Sutton does not teach any of the 3 operations while the Office asserted previously that Sutton taught something about the altering. In any case, Applicant agrees with the Office that Sutton does not teach that "the text data [is] processed into converted text data."

In addition, the Office has asserted that Dietz teaches "the text data being processed into converted text data" in Figure 2; column 5, line 40 to column 6, line 27; Figure 3; and column 6, lines 24-67. Dietz teaches the following:

“When the text is accurate, the process then implements machine language conversion software to convert text in L₁ to text in language 2 (L₂) (step 319)” (col. 6, lines 51–53, emphasis added).

Therefore, Dietz teaches that the text to converted-text processing (operation 2) includes language translation. Since Sutton does not teach operation 2, as previously discussed, the combination of Sutton and Dietz must include language translation in operation 2. Regardless of which reference teaches operation 3 (synthesizing the converted text data into audio data that includes the applied expression that does not perform language translation), any resulting audio data or applied expression will include language translation because the synthesizing is based in the converted text data, and the converted text data includes language translation. For all these reasons, the combination of Sutton and Dietz does not teach the aforementioned claim because the combination will always perform language translation.

Applicant notes that the Office is trying a combination by choosing operations from Sutton and Dietz that are not compatible. The Office has asserted that altering the content data is taught by both references, but it is not clear how this combination could even operate properly. Further, the Office seems to choose operations from Sutton that are not compatible with Dietz. For example, Sutton operates on voice data while Dietz operates on text to perform the language translation. The Office must select operations from one or the other, such that the claimed result is obtained. Selecting “does not perform language translation” from Sutton, while “selecting converting text to text data” from Dietz would not create a result that does not perform language translation, as discussed above.

Further, Applicant notes that Dietz was previously used and then withdrawn as a reference because it was determined that Dietz teaches language translation. Applicant notes

that having to repeat previously presented arguments generates an unnecessary delay in the prosecution of the Patent Application.

2. Combining Dietz with Sutton would change the principle of operation of Sutton

Sutton teaches the following:

"A method of producing synthetic visual speech according to this invention includes receiving an input containing speech information. One or more visemes that correspond to the speech input are then identified. Next, the weights of those visemes are calculated using a coarticulation engine including viseme deformability information. Finally, a synthetic visual speech output is produced based on the visemes' weights over time (or tracks). The synthetic visual speech output is combined with a synchronized audio output corresponding to the input to produce a multimedia output containing a 3D lipsyncing animation" (Abstract, emphasis added);

"...a viseme is a visual speech representation defined by the external appearance of articulators (i.e., lips, tongue, teeth, etc.) during articulation of a corresponding phoneme" (col. 1, lines 17-21); and

"According to this process 1A, a user inputs a voice file 2B and a text file 2A representing the same speech input into the system. The text file 2A must correspond exactly to the voice file 2B in order for the process to work properly. The system 1A then takes the voice and text inputs 2B, 2A and forces an alignment between them in a forced alignment generator 18. Because the text input 2A informs the system 1A of what the voice input 2B says, there is no need to attempt to separately recognize the phonetic components of the speech input from the voice file 2B, for example, by using a speech recognition program" (col. 16, lines 12-23, emphasis added).

Sutton teaches producing synthetic visual speech based on visemes, which are visual speech representations defined by the external appearance of articulators during articulation of a corresponding phoneme. Therefore, Sutton is concerned with articulation of phonemes, and not with the actual content of the speech.

On the other hand, Dietz teaches the following:

"A method for dynamically providing language translations of a human utterance from a first human language into a second human language. A human utterance is captured in a first human language utilizing a speech input device. The speech input device is then linked to a server created from components including a data processing system

equipped with software enabled speech recognition environment and a language translation environment. A desired second human language is then selected for said first human language to be translated into. Following this selection, the captured human utterance is transmitted to the server where it is converted into text utilizing the speech recognition engine of the server which instantiates the translation of the text from the first human language into the desired second human language. Finally, an output is provided of the captured human utterance in its desired second human language" (Abstract, emphasis added).

Dietz teaches to provide language translation of human utterances. Sutton teaches that "there is no need to ... [use] a speech recognition program." However, Dietz does teach a speech recognition program. Since Dietz indicates that a speech recognition is not needed, using speech recognition would alter the principle of operation.

Further, Sutton teaches that "a user inputs a voice file 2B and a text file 2A representing the same speech input into the system ... [that] must correspond exactly." If Dietz is combined with Sutton, text translation would take place, and the text data would no longer correspond with the voice file. As a result, the visual speech created would not match the audio file (i.e., the lips of the speaker would not be in sync with the voice). Also, the person skilled in the art would have no motivation to make the combination because the combination would not work. For these reasons, a combination of Sutton and Dietz would not operate properly or the combination would change the principle of operation of Sutton.

3. The Office has not provided articulated reasoning with rational underpinning to support the legal conclusion of obviousness

The Office has asserted that "[i]t would have been obvious ... to incorporate the teaching of Dietz ... in the method of Sutton ... because it would have increased the round-trip processing speed and provided the system for providing synthesized audio data to improve

speech communication between two computers” (page 4, 2nd para.) Applicant respectfully disagrees. There is no rational underpinning to the reason provided by the Office.

The Supreme Court in KSR noted that the analysis supporting a rejection under 35 U.S.C. 103 should be made explicit, *In re KSR International Co. v. Teleflex Inc. (KSR)*, 550 U.S. ___, 82 USPQ2d 1385 (2007). The Court in KSR quoted *In re Kahn*, which stated that “[R]ejections on obviousness cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.”

The Office has merely put forth conclusory statements and not provided articulated reasoning to support the legal conclusion of obviousness. The Office has not explained how the combination would increase processing speed or how it would improve speech communication.

Further, asserting that “it would have increased the round-trip processing speed” is not rational. Adding Dietz to Sutton would mean converting voice to text, translating text, and then converting to voice again. It is not possible to increase speed to a method by adding additional steps, such as converting to text, translating, etc. Therefore, the reason articulated by the Office has no rational underpinning to support the legal conclusion of obviousness.

4. Conclusion

Independent claims 10, 14, 22, 30, 32, and 37 are believed to be patentable for at least the same reasons that claim 1 is believed to be patentable. In view of the foregoing, the Office is requested to withdraw the rejection of claims 1, 10, 14, 22, 30, 32, and 37 under §103. The dependent claims are submitted to be patentable for at least the same reasons that the independent claims are believed to be patentable. The Applicants therefore respectfully

request reconsideration and allowance of the pending claims. A Notice of Allowance is respectfully requested.

If the Examiner has any questions concerning the present amendment, the Examiner is kindly requested to contact the undersigned at (408) 774-6903. If any other fees are due in connection with filing this amendment, the Commissioner is also authorized to charge Deposit Account No. 50-0805 (Order No. SONYP009).

Respectfully submitted,
MARTINE PENILLA & GENCARELLA, LLP

/Albert Penilla/

Albert S. Penilla, Esq.
Reg. No. 39,487

710 Lakeway Drive, Suite 200
Sunnyvale, CA 94085
Telephone: (408) 774-6920
Facsimile: (408) 749-6901
email: jose@mpiplaw.com
Customer Number 25920